

Data Back-Up and Recovery Techniques for Cloud Server Using Seed Block Algorithm

R. V. Gandhi¹, M Sessaiah², A. Srinivas³, C. ReddiNeelima⁴

Department of Computer Science and Engineering, JigjigaUnivesity, Jigjiga, Ethiopia.

Department of Computer Science and Engineering, Visvesvaraya Technological University, Belgum, India.

Department of Computer Science and Engineering, HITS, R.R.Dist, Telangana, India.

Department of Computer Science and Engineering, MTIEAT, Palamaner, Chittoor Dist, A.P, India.

Abstract

In cloud computing, data generated in electronic form are large in amount. To maintain this data efficiently, there is a necessity of data recovery services. To cater this, we propose a smart remote data backup algorithm, Seed Block Algorithm. The objective of proposed algorithm is twofold; first it help the users to collect information from any remote location in the absence of network connectivity and second to recover the files in case of the file deletion or if the cloud gets destroyed due to any reason. The time related issues are also being solved by proposed seed block algorithm such that it will take minimum time for the recovery process. Proposed seed block algorithm also focuses on the security concept for the back-up files stored at remote server, without using any of the existing encryption techniques.

Key Words: Data Recovery, Seed Block Algorithm, Security, center repository, backup repository.

I. INTRODUCTION

Cloud computing defines as a model for enabling convenient, on-demand network access to a share pool of configurable computing service that can be provisioned rapidly and released with minimal management effort or services provider. Today, Cloud Computing is itself a gigantic technology which is surpassing all the previous technology of computing of this competitive and challenging IT world. The need of cloud computing is increasing day by day as its advantages overcome the disadvantage of various early computing techniques. Cloud storage provides online storage where data stored in form of virtualized pool that is usually hosted by third parties. The hosting company operates large data on large data centre and according to the requirements of the customer these data centre virtualized the resources and expose them as the storage pools that help user to store files or data objects. As number of user shares the storage and other resources, it is possible that other customers can access your data. Either the human error, faulty equipment's, network connectivity, a bug or any criminal intent may put our cloud storage on the risk and danger. And changes in the cloud are also made very frequently; we can term it as data dynamics. The data dynamics is supported by various operations such as insertion, deletion and block modification. Since services are not limited for archiving and taking backup of data; remote data integrity is also needed. Because the data integrity always focuses on the validity and fidelity of the complete state of the server that

takes care of the heavily generated data which remains unchanged during storing at main cloud remote server and transmission. Integrity plays an important role in back-up and recovery services. However, still various successful techniques are lagging behind some critical issues like implementation complexity, low cost, security and time related issues. To cater these issues, we propose a smart remote data backup algorithm, Seed Block Algorithm (SBA). The contribution of the proposed SBA is twofold; first SBA helps the users to collect information from any remote location in the absence of network connectivity and second to recover the files in case of the file deletion or if the cloud gets destroyed due to any reason.

II. RELATED WORK

2.1. Problem statement

In cloud computing, to maintain the data efficiently, there is a necessity of data recovery services. To cater this, we propose a smart remote data backup algorithm, Seed Block Algorithm. Using Seed block algorithm we recover the files in case of the file deletion or if the cloud gets destroyed due to any reason. The time related issues are also being solved by proposed Seed Block Algorithm such that it will take minimum time for the recovery process. Proposed Seed Block Algorithm also focuses on the security concept for the back-up files stored at remote server, without using any of the existing encryption techniques.

2.2. Existing Statement

The recent back-up and recovery techniques that have been developed in cloud computing domain such are HSDRT, (PCS) Parity Cloud Service, (ERGOT) Efficient Routing Grounded on Taxonomy, Linux Box, Cold/Hot backup strategy etc. Detail review shows that none of these techniques are able to provide best performances under all uncontrolled circumstances such as cost, security, low implementation complexity, redundancy and recovery in short span of time.

2.3. Proposed System

The objective of proposed algorithm is twofold; first it help the users to collect information from any remote location in the absence of network connectivity and second to recover the files in case of the file deletion or if the cloud gets destroyed due to any reason.

Usually Backup server of main cloud is the copy of main cloud. When this Backup Server is at remote location (i.e. far away from the main server) and having the complete state of the main cloud, and then this remote location server is termed as Remote Data Backup Server. The main cloud is termed as the central repository and remote backup cloud is termed as remote repository. To tackle the challenges like low implementation complexity, low cost, security and time related issues we propose Seed Block Algorithm (SBA) algorithm.

III. IMPLEMENTATION

These aim to objectively and rationally uncover the strengths and weaknesses of the existing business or proposed venture, opportunities and threats as presented by the environment, the resources required to carry through, and ultimately the prospects for success. In its simplest term, the two criteria to judge feasibility are cost required and value to be attained. As such, a well-designed feasibility study should provide a historical background of the business or project, description of the product or service, accounting statements, details of the operations and management, marketing research and policies, financial data, legal requirements and tax obligations. Generally, feasibility studies precede technical development and project implementation.

Cloud Server - Central Repository

In some respects cloud servers work in the same way as physical servers but the functions they provide can be very different. When opting for cloud hosting, clients are renting virtual server space rather than renting or purchasing physical servers. They are often paid for by the hour depending on the capacity required at any particular time. Traditionally there are two main options for hosting: shared hosting and dedicated hosting. Shared hosting is the cheaper option whereby servers are shared between the hosting provider's clients. One client's website will be hosted on the same server as websites belonging to other clients. This has several disadvantages including the fact that the setup is inflexible and cannot

cope with a large amount of traffic. Dedicated hosting is a much more advanced form of hosting, whereby clients purchase whole physical servers. This means that the entire server is dedicated to them with no other clients sharing it. In some instances the client may utilize multiple servers which are all dedicated to their use. Dedicated servers allow for full control over hosting. The downside is that the required capacity needs to be predicted, with enough resource and processing power to cope with expected traffic levels. If this is underestimated then it can lead to a lack of necessary resource during busy periods, while overestimating it will mean paying for unnecessary capacity. With cloud hosting clients get the best of both worlds. Resource can be scaled up or scaled down accordingly, making it more flexible and, therefore, more cost-effective. When there is more demand placed on the servers, capacity can be automatically increased to match that demand without this needing to be paid for on a permanent basis. This is akin to a heating bill; you access what you need, when you need it, and then only pay for what you've used afterwards. Unlike dedicated servers, cloud servers can be run on a hypervisor. The role of a hypervisor is to control the capacity of operating systems so it is allocated where needed. With cloud hosting there are multiple cloud servers which are available to each particular client. This allows computing resource to be dedicated to a particular client if and when it is necessary. Where there is a spike in traffic, additional capacity will be temporarily accessed by a website, for example, until it is no longer required. Cloud servers also offer more redundancy. If one server fails, others will take its place. Cloud computing is the provision of dynamically scalable and often virtualized resources as a services over the internet Users need not have knowledge of, expertise in, or control over the technology infrastructure in the "cloud" that supports them. Cloud computing represents a major change in how we store information and run applications. Instead of hosting apps and data on an individual desktop computer, everything is hosted in the "cloud server" called the central repository an assemblage of computers and servers accessed via the Internet. User can use the cloud server for storing the data in a secure manner. For using the cloud server; user can register first for getting the user id and password. If the user is already registered then user can use the cloud server with the user id and password.

Backup Repository

The main cloud is termed as the central repository and remote backup cloud is termed as Backup repository. And if the central repository lost its data under any circumstances either of any natural calamity (for ex - earthquake, flood, fire etc.) or by human attack or deletion that has been done mistakenly and then it uses the information from the remote repository. The main objective of the backup facility is to help user to collect information from any remote location even if network

connectivity is not available or if data not found on main cloud. Backups have two distinct purposes. The primary purpose is to recover data after its loss, be it by data deletion or corruption. Data loss can be a common experience of computer users. The secondary purpose of backups is to recover data from an earlier time, according to a user-defined data retention policy, typically configured within a backup application for how long copies of data are required. Though backups popularly represent a simple form of disaster recovery, and should be part of a disaster recovery plan, by themselves, backups should not alone be considered disaster recovery. One reason for this is that not all backup systems or backup applications are able to reconstitute a computer system or other complex configurations such as a computer cluster, directory servers, or a database server, by restoring only data from a backup.

Since a backup system contains at least one copy of all data worth saving, the data storage requirements can be significant. Organizing this storage space and managing the backup process can be a complicated undertaking. A data repository model can be used to provide structure to the storage. Nowadays, there are many different types of data storage devices that are useful for making backups. The Remote backup services should cover the following issues:

- Privacy and ownership.
- Relocation of servers to the cloud.
- Data security.
- Reliability.
- Cost effectiveness.

Privacy and ownership

Different clients access the cloud with their different login or after any authentication process. They are freely allowed to upload their private and essential data on the cloud. Hence, the privacy and ownership of data should be maintained; Owner of the data should only be able to access his private data and perform read, write or any other operation. Remote Server must maintain this Privacy and ownership.

Relocation of server

For data recovery there must be relocation of server to the cloud. The Relocation of server means to transfer main server's data to another server; however the new of location is unknown to the client. The clients get the data in same way as before without any intimation of relocation of main server, such that it provides the location transparency of relocated server to the clients and other third party while data is been shifted to remote server.

Data security

The client's data is stored at central repository with complete protection. Such a security should be followed in its remote repository as well. In remote

repository, the data should be fully protected such that no access and harm can be made to the remote cloud's data either intentionally or unintentionally by third party or any other client.

Reliability

The remote cloud must possess the reliability characteristics. Because in cloud computing the main cloud stores the complete data and each client is dependent on the main cloud for each and every little amount of data; therefore the cloud and remote backup cloud must play a trustworthy role. That means, both the server must be able to provide the data to the client immediately whenever they required either from main cloud or remote server.

Cost effectiveness

The cost for implementation of remote server and its recovery & back-up technique also play an important role while creating the structure for main cloud and its correspondent remote cloud. The cost for establishing the remote setup and for implementing its technique must be minimum such that small business can afford such system and large business can spend minimum cost as possible.

IV. RESULTS



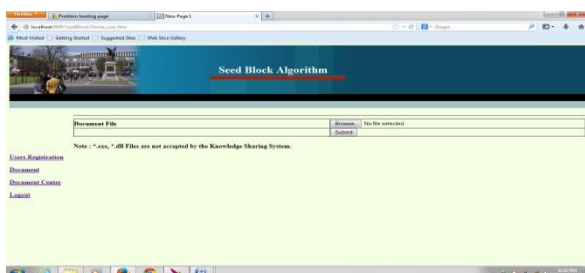
Screen Shot For Login Page



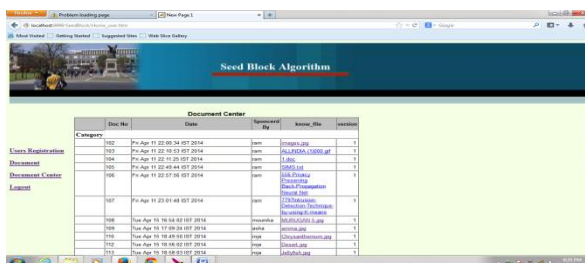
Screen shot for module list



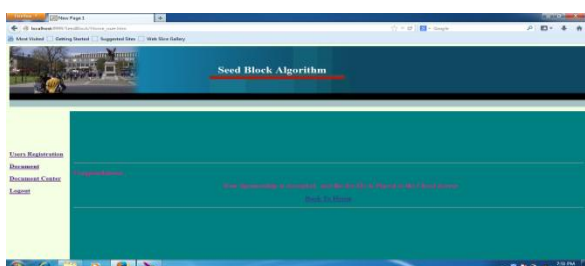
Screen shot for User Registration form



Screen shot for browsing a document



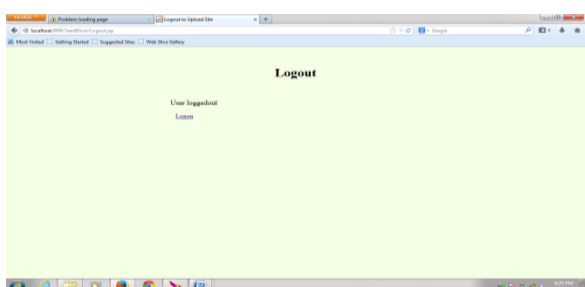
Screen shot for Document center



Screen shot for Document Acceptance



Screen shot displaying File Corruption



Screen Shot For Logout

V. CONCLUSION

In this project, we presented detail design of proposed SBA (seed Block algorithm) algorithm. Proposed SBA is robust in helping the users to collect

information from any remote location in the absence of network connectivity and also to recover the files in case of the file deletion or if the cloud gets destroyed due to any reason. Experimentation and result analysis shows that proposed SBA also focuses on the security concept for the back-up files stored at remote server, without using any of the existing encryption techniques. The time related issues are being solved by proposed SBA such that it will take minimum time for the recovery process. Many reputed companies can store the confidential data without much risk, Security is also being increased.

REFERENCES

- [1] Yoichiro Ueno, Noriharu Miyaho, Shuichi Suzuki, Muzai Gakuendai, Inzai-shi, Chiba, Kazuo Ichihara, 2010, "Performance Evaluation of a Disaster Recovery System and Practical Network System Applications," Fifth International Conference on Systems and Networks Communications, pp 256-259.
- [2] Chi-won Song, Sungmin Park, Dong-wook Kim, Sooyong Kang, 2011, "Parity Cloud Service: A Privacy-Protected Personal Data Recovery Service," International Joint Conference of IEEE TrustCom-11/IEEE ICESS-11/FCST-11.
- [3] Y.Ueno, N.Miyaho, and S.Suzuki, "Disaster Recovery Mechanism using Widely Distributed Networking and Secure Metadata Handling Technology", Proceedings of the 4th edition of the UPGRADE-CN workshop, 2009, pp. 45-48.
- [4] Giuseppe Pirr'o, Paolo Trunfio, Domenico Talia, Paolo Missier and Carole Goble, "ERGOT: A Semantic-based System for Service Discovery in Distributed Infrastructures," 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, 2010.
- [5] Vijaykumar Javaraiah Brocade Advanced Networks and Telecommunication Systems (ANTS), "Backup for Cloud and Disaster Recovery for Consumers and SMBs," IEEE 5th International Conference, 2011.
- [6] Lili Sun, Jianwei An, Yang Yang, Ming Zeng, "Recovery Strategies for Service Composition in Dynamic Network," International Conference on Cloud and Service Computing, 2011.
- [7] Xi Zhou, Junshuai Shi, Yingxiao Xu, Yinsheng Li and Weiwei Sun, 2008, "A backup restoration algorithm of service composition in MANETs," Communication Technology ICCT 11th IEEE International Conference, pp. 588-591.
- [8] M. Armbrust et al, "Above the clouds: A Berkeley view of cloud computing," <http://www.eecs.berkeley.edu/>, 2009.

- [9] F.BKashani, C.Chen, C.Shahabi.WSPDS, 2004, "Web Services Peer-to-Peer Discovery Service," ICOMP.
- [10] Eleni Palkopoulou, Dominic A. Schupke, Thomas Bauschert, 2011, "Recovery Time Analysis for the Shared Backup Router Resources (SBRR) Architecture", IEEE ICC.
- [11] Balazs Gerofi, Zoltan Vass and Yutaka Ishikawa, "Utilizing Memory Content Similarity for Improving the Performance of Replicated Virtual Machines", Fourth IEEE International Conference on Utility and Cloud Computing 2011.
- [12] P.Demeester et al., 1999. , "Resilience in Multilayer Networks," IEEE Communications Magazine, Vol. 37, No. 8, p.70-76. S. Zhang, X. Chen, and X. Huo, 2010, "Cloud Computing Research and Development Trend," IEEE Second International Conference on Future Networks, pp. 93-97.
- [13] T. M. Coughlin and S. L. Linfoot, 2010, "A Novel Taxonomy for Consumer Metadata," IEEE ICCE Conference.
- [14] Kruti Sharma, Kavita R Singh "Online Data Back-up and Disaster Recovery Techniques in Cloud Computing: A Review", IJEIT, Volume 2, Issue 5, November 2012.

engineering (B.Tech) from VTU, Belgum, 2003. He is working as an Assistant Professor in Department of Computer Science & Engineering in **S.J.C. Institute of Technology, Chickballapur**. He has 11+ years of Teaching Experience. His Research Interests Include Network Security, Cloud Computing & Data Warehousing and Data Mining.



A.Srinivas, Post Graduated in Computer Science & Engineering (**M.Tech**) From JNT **University**, Hyderabad in 2009 and Graduated in Computer Science & Information Technology (B.Tech) from **JNTU**, Hyderabad in 2004. He is currently working as an Assistant Professor, Department of Computer Science & Engineering in **Holy Mary Institute of Technology and Science (HITSCOE)**, (V) Bogaram, (M) Keesara, R.R. Dist, Telangana, India. He has 8+ years of Teaching Experience. His research interests include Cloud Computing, Data Mining, Information Security, Software Testing, Wireless Networks and Software Quality.

AUTHOR'S PROFILE



Mr. **R.VENKATESWARA GANDHI**, Post Graduated in Computer Science & Engineering (**M.Tech**) From JNT **University**, Hyderabad in 2009 and Graduated in Computer Science & engineering (B.Tech) form **JNTU**, Hyderabad, 2007. He is working as an Lecturer in Department of Computer Science & Engineering in **Jigjiga Univesity, Jigjiga, Ethiopia**. He has 5+ years of Teaching Experience. His Research Interests Include Network Security, Cloud Computing & Data Warehousing and Data Mining.



Mr. **M Sessaiah**, Post Graduated in Computer Science & Engineering (**M.Tech**) From **Visvesvaraya Technological University**, Belgum in 2010 and Graduated in Computer Science &



Mrs C.Reddi Neelima, Post Graduated in Computer Science & Engineering (M.Tech), JNTU University, Anantapur in 2014 and Graduated in Computer Science Engineering (B.Tech) from SK university, Anantapur in 2008. She is working as an Assistant Professor in Department of Computer Science & Engineering in **Mother Theresa Institute of Engineering and Technology**, Chittoor Dist, A.P and India. She has 3 years of Teaching Experience. Her Research Interests Include Network Security, Cloud Computing & Data Warehousing and Data Mining.